Incorporating advanced behavioral models in mixed linear optimization

Meritxell Pacheco
Shadi Sharif Azadeh, Michel Bierlaire

Transport and Mobility Laboratory (TRANSP-OR)
École Polytechnique Fédérale de Lausanne

June 14th, 2016

Outline

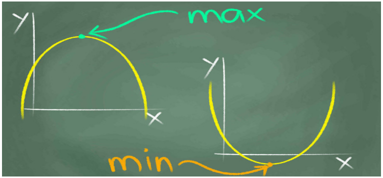
Motivation

Customer behavioral models

Demand

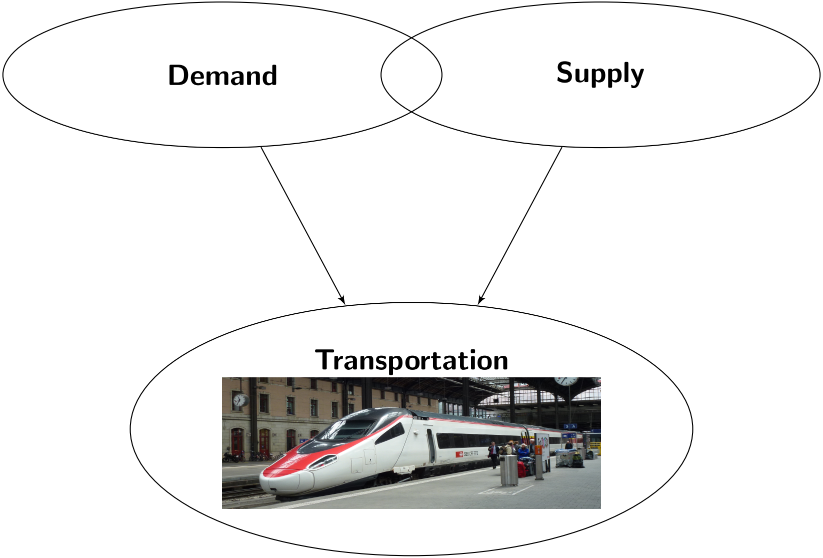
Motivation

Operations research



Supply

Motivation



Demand and supply

Customer behavioral models

- **Given**: configuration of the system ⇒ predict the demand
- Maximize satisfaction
- **Here:** discrete choice models

Operations Research

- **Given**: demand ⇒ configure the system
- Minimize costs
- **Here:** MILP

Discrete choice models in optimization problems

- Integration of choice models ⇒ source of non convexity
- Many techniques to convexify and linearize
- **Here**: different approach that addresses
  - Nonconvex representation of choice probabilities
  - Wide class of discrete choice models

Utilities

Demand and supply

- Population of N individuals
- Set of products $\mathcal{C}$ in the market
  - artificial "opt-out" product
- $\mathcal{C}_n \subseteq \mathcal{C}$ subset of available products to individual n

Utility

$U_{in} = V_{in} + \varepsilon_{in}$: associated score with alternative i by individual n

- V_{in}: deterministic part
- ε_{in}: error term

Behavioral assumption: n chooses i if U_{in} is the highest in $\mathcal{C}_n$

Probabilistic model

Choice

$$w_{in} = \begin{cases} 1 & \text{if } n \text{ chooses } i \\ 0 & \text{otherwise} \end{cases}$$

$\forall n, \forall i \in \mathcal{C}$

Availability

$$y_{in} = \begin{cases} 1 & \text{if } i \in \mathcal{C}_n \\ 0 & \text{otherwise} \end{cases}$$

$\forall n, \forall i \in \mathcal{C}$

Probabilistic model

- $\Pr(w_{in} = 1) = \Pr(U_{in} \geq U_{jn}, \forall j \in \mathcal{C}_n)$ and i available ($y_{in} = 1$)
- $D_i = \sum_{n=1}^{N} \Pr(w_{in} = 1)$
- D_i is in general non linear
- **Example:** $\Pr(w_{in} = 1) = \frac{y_{in} e^{V_{in}}}{\sum_{j \in \mathcal{C}} y_{jn} e^{V_{jn}}}$ (logit model)

Simulation

Simulation

- Assume a distribution for ε_{in}
- Generate R draws $\xi_{in1} \dots \xi_{inR}$
- r behavioral scenario
- The choice problem becomes **deterministic**

Demand model

$$U_{inr} = V_{in} + \xi_{inr} = \sum_k \beta_k x_{ink} + f(z_{in}) + \xi_{inr} \Rightarrow \text{not a random variable} \quad (1)$$

- **Endogeneous** part of V_{in}: x_{ink} decision variables, linear (assumption)
- **Exogeneous** part of V_{in}: other variables z_{in}, f not necessarily linear

Availability of alternatives

Variables

- y_{in} decision of the operator

$$y_{in} = 0 \qquad \forall i \notin \mathcal{C}_n, n \tag{2}$$

- y_{inr} availability at scenario level (*e.g.* demand exceeding capacity)

$$y_{inr} \leq y_{in} \qquad \forall i, n, r \tag{3}$$

Idea: auxiliar variable to consider only the utilities of the available alternatives

$$\nu_{inr} = \begin{cases} U_{inr} & \text{if } y_{inr} = 1 \\ \ell_{nr} & \text{if } y_{inr} = 0 \end{cases}$$

where $\ell_{nr} = \min_{j \in \mathcal{C}_n}\{U_{jnr}\}$

Highest utility and choice

Linearization of the maximum of variables

$U_{nr} = \max_{j \in \mathcal{C}_n} \{U_{jnr}\}$

Highest utility for individual n in scenario r: $\mu_{inr} = \begin{cases} 1 & \text{if } U_{nr} = U_{inr} \\ 0 & \text{otherwise} \end{cases}$

Highest utility, choice and availability

w_{inr} choice variable at scenario level

- An unavailable alternative cannot be the one with highest utility
- An alternative without the highest utility cannot be chosen
- Only one alternative is chosen

Modeling framework

Summary

- Introduced model is linear in...
  - Any variable appearing linearly in U_{inr}
  - The availability variables y_{in}, y_{inr} and ν_{inr}
  - The preference variables μ_{inr}
  - The choice variables w_{inr}
- Demand within the market

$$D_i = \frac{1}{R} \sum_{n=1}^{N} \sum_{r=1}^{R} w_{inr}$$

- Further specifications
  - Capacity?
  - Price?

Maximization of revenues

Application

- Operator selling services to a market, each service:
  - Price
  - Capacity (number of customers)
- Demand is price elastic and heterogenous
- **Goal:** best strategy in terms of capacity allocation and pricing

Revenues

- p_{in} price that individual n has to pay to access to service i

$$R_i = \frac{1}{R} \sum_{n=1}^{N} p_{in} \sum_{r=1}^{R} w_{inr}$$

- p_{in} endogenous variable $\Rightarrow$ R_i non linear

Pricing

Linearization of R_i

- Discretization of the price $\Rightarrow p_{in}^1, \ldots, p_{in}^{L_{in}}$
- Binary variables λ_{inl} such that $p_{in} = \sum_{l=1}^{L_{in}} \lambda_{inl} p_{in}^l$ and $\sum_{l=1}^{L_{in}} \lambda_{inl} = 1$
- Revenues for alternative i

$$R_i = \frac{1}{R} \sum_{n=1}^{N} \sum_{l=1}^{L_{in}} \lambda_{inl} p_{in}^l \sum_{r=1}^{R} w_{inr}$$

- Still non linear $\Rightarrow \alpha_{inrl} = \lambda_{inl} w_{inr}$ to linearize it
- **Objective function**

$$\max R_i = \max \frac{1}{R} \sum_{n=1}^{N} \sum_{l=1}^{L_{in}} \alpha_{inrl} p_{in}^l$$

Capacity

Overview

- c_i capacity of service i
- Who has access if the capacity is reached?
- The model favors customers bringing higher revenues
- ... but generally customers arrive in a random order

Priority list

- An individual is served only if all individuals before her in the list have been served
- $y_{inr} \geq y_{i(n+1)r} \ \forall i, n, r$
- Can account for fidelity programs, VIP customers, etc.
- We assume it given

Parking choices

Original experiment

- [Ibeas et al., 2014] *Modelling parking choices considering user heterogeneity*
- Stated preferences survey (197 respondents, 8 scenarios)
- Analyze viability of an underground car park

Free on-Street Parking (FSP)

Free

Paid on-Street Parking (PSP)

Price levels: 0.6 and 0.8

Paid Underground Parking (PUP)

Price levels: 0.8 and 1.5

Choice model and preliminary experiments

Mixed Logit model

- **Attributes:** time to reach the destination
- **Random parameters:** access time and price
- **Socioeconomic characteristics:** residence, age of the vehicle
- **Interactions:** price and low income, price and residence

	Low Income	**Not low income**
Resident	39.09 %	14.21 %
Non resident	30.96 %	15.74 %

	AT	**TD**	**FEE**
FSP	15	10	0
PSP	10	15	0.6
PUP	5	10	1.5

Preliminary experiments

Assumptions

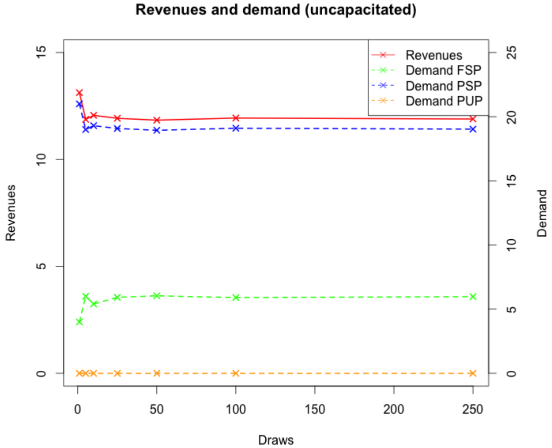
- Subset of 25 individuals
- FSP as opt-out alternative
- Capacity defined to challenge the algorithm (uncapacitated for FSP and 8 individuals for PSP and PUP)

Price levels

- Set a price lower and upper bound for each alternative
- Divide the interval into 5 equidistant price levels
- Run it for $R = 100$ draws
- Generate a reduced the interval around the obtained solution until no improvement is obtained

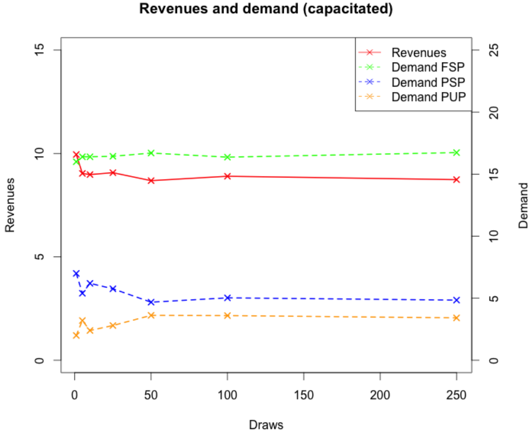
Price levels (uncapacitated)

- **PSP**: 0.31, 0.47, 0.63, 0.78, 0.94
- **PUP**: 0.31, 0.47, 0.63, 0.78, 0.94

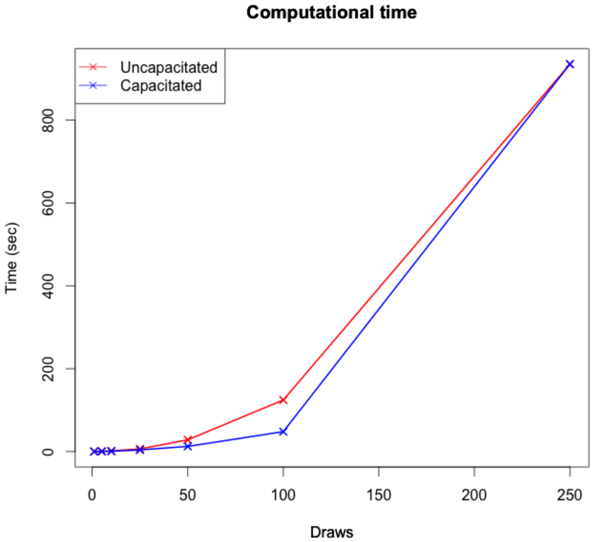


Price levels (capacitated)

- **PSP**: 1.10, 1.13, 1.15, 1.18, 1.20
- **PUP**: 0.90, 0.91, 0.93, 0.94, 0.95



Computational time



Conclusions and future work

Conclusions

- High dimensionality of the problem
- Price levels calculation
- Any assumption can be made for the ε_{in}

Future work

- Design of scenarios $\Rightarrow$ more experiments!
- Speed up the computational results: decomposition techniques
    - By **customer**: capacity!
    - By **scenario**: only considered together in the objective function
- Introduce new features (*e.g.* N as a group of individuals), capacity?

Questions?

Bibliography

A. Ibeas, L. dell'Olio, M. Bordagaray, and J. de D. Ortúzar. Modelling parking choices considering user heterogeneity. *Transportation Research Part A: Policy and Practice*, 70:41 – 49, 2014. ISSN 0965-8564. doi: http://dx.doi.org/10.1016/j.tra.2014.10.001. URL http://www.sciencedirect.com/science/article/pii/S0965856414002341.